Refining Statewide School Accountability Systems

Lessons From Innovative State Approaches Under ESSA

Acknowledgment

Introduction

During the past 2 decades, federal education policy has aimed to identify lower performing schools and focus school improvement efforts through state-level school accountability systems. However, throughout the lifespans of the No Child Left Behind Act and the Every Student Succeeds Act (ESSA), the connection between the identification of schools most in need of support and school improvement on valued outcomes often has been limited (Carlson le Floch & Atchison, 2025; Munyan-Penney et al., 2024; Pinsonneault, 2023). As the country has emerged from the COVID-19 pandemic, states have increasingly acknowledged the need to expand accountability beyond what is required by ESSA. A handful of states have participated in the U.S. Department of Education's Innovative Assessment Demonstration Authority (IADA), which seeks to establish an innovative assessment system that can be used for accountability purposes. However, many interest holders in the education field have identified shortcomings of the IADA, including a lack of support in developing pilot projects and regulatory constraints that limit innovation (EdTrust, 2023). In addition, multiple states have withdrawn from the program.

To address the challenges educators face with the ESSA and IADA accountability policies, some states have built stand-alone accountability systems (Burnette, 2018) that operate in addition to the federally mandated system. Other states have poured resources into developing metrics beyond what is required by ESSA (e.g., military readiness in states like Ohio and North Dakota, school climate measures in states like California and Illinois). Concurrently, accountability policy changes being discussed and enacted at both the state and federal levels will have large implications for accountability practices across the country. Many states (e.g., California, Massachusetts, and Vermont) have either made or are considering changes to their state ESSA plans so that they better reflect the postpandemic context.

Meanwhile, changes in federal education policy under the second Trump Administration and the prospect of the Elementary and Secondary Education Act (ESEA) reauthorization loom over state-level accountability policy discussions.

Although no provisions of ESSA have been changed under the current administration, questions remain about how strongly state accountability requirements will be enforced and whether additional flexibilities will be granted. In addition, it is unclear when (or if) the U.S. Congress will take up ESEA reauthorization, but past reauthorizations have fundamentally altered the accountability landscape.

Responses to the current policy context that are guided by research-based recommendations taken from current state-level practices offer the promise of better informing both state-level actors who implement accountability systems and federal policymakers as they consider future changes to education statutes. However, little research has been conducted on innovative accountability approaches that fall outside the bounds of ESSA. Much of the research that has been conducted on accountability systems under ESSA has focused on how states are complying with ESSA requirements (e.g., the comparison of state school accountability systems by the Education Commission of the States). Given the rapid expansion of innovative state approaches to accountability, research examining these approaches is needed to both understand the current state of accountability practices and guide future policymaking.

Objectives

This white paper had two main objectives. First, it documents approaches that states have taken to expand the concept and implementation of school accountability in order to better understand the current context of state accountability systems. Since ESSA was enacted in late 2015, many systems have shifted, sometimes dramatically. Although there was a surge of focus on accountability systems as states were submitting their first ESSA plans, relatively little research has followed the innovation that has occurred since, particularly following the pandemic. This paper provides resources for state-level leaders, policymakers, and other interest holders about innovative approaches that have been attempted and the strengths and challenges of these approaches.

Second, this white paper provides relevant information about potential approaches to future change in statewide school accountability systems at the state and federal levels in order to better connect school improvement efforts and substantive improvements in student outcomes. To that end, this paper surfaces important concepts to consider as state leaders undertake adjustments to statewide school accountability systems in order to elevate promising state-level policies and practices. By extension, this paper also provides important topics for federal policymakers to consider as they approach the potential reauthorization of ESEA, with the goal of encouraging connections between school accountability and improvement.

Research Objectives

To provide richer detail about innovative practices implemented by state and local educational agencies (SEAs and LEAs), the project team identified four case studies that provide deeper insights about approaches to innovation that could guide other state accountability systems and inform future accountability policy discussions. In selecting these case studies, the project team focused on three types of innovations:

- use of innovative accountability indicators, both in terms of unique school quality and success indicators and innovative approaches to other indicators

- use of dual accountability systems, which involve state implementation of additional accountability practices outside the federal requirements under ESSA

- implementation by LEAs or other organizations of accountability practices that introduce more flexibility at the district and school levels

Research Questions

To further refine the approach to the case studies examined in this project, the project team defined four main research questions that each case study will address:

1. What is the context of states that have embraced innovative accountability practices? What brought about the innovations, and how have they been used?

2. What impact have innovative accountability practices had within the state?

3. What successes and challenges have SEAs and LEAs faced in implementing innovative accountability practices?

4. How can the experiences of implementing these innovative accountability practices inform federal and state policy?

Guiding Principles for the Research

To ensure that the case studies provided relevant and actionable information for multiple audiences, the project team also identified three guiding principles for them. These principles provided the underlying foundation for the research and assisted the project team in establishing a common approach and tone for this work.

- The experiences of the staff of the organizations that implement innovative accountability practices should be the focus of the case studies. The information we learn from these accountability experts should guide the policy recommendations found in this research.

- The innovations highlighted in the case studies should range in size and scale of implementation from small to large to allow states and organizations to choose innovations that match the capacity and appetite for change.

- The case studies should focus on the processes used to implement accountability innovations so that interested organizations understand the ins and outs of implementing these types of changes. For example, focusing on outcomes at inappropriate times (i.e., expecting student improvement on assessments before the innovative practice has had time to mature) can distract from the quality of implementation needed to produce measurable outcomes.
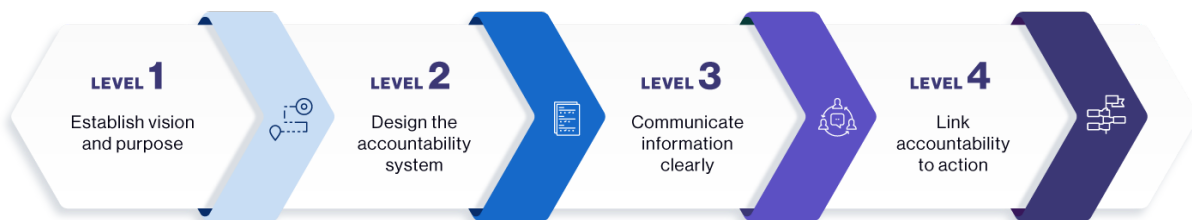
The Role of Statewide School Accountability Systems

As demonstrated both in the case studies presented in this paper and in practical applications, school accountability systems serve many different purposes within the education ecosystem. For example, the National Center for the Improvement of Educational Assessment contends that statewide school accountability systems play four key roles:

- building public trust and engagement

- signaling what is important to district and school leaders

- monitoring group, school, and district performance

- supporting school improvement initiatives

Based on the findings from the case studies, we propose a model of statewide school accountability systems that expands on these ideas, refining the roles and creating a hierarchical relationship between the roles (Figure 1).

Figure 1. Model of Statewide School Accountability Systems

LEVEL **1** Establish vision and purpose

LEVEL **2** Design the accountability system

LEVEL **3** Communicate information clearly

LEVEL **4** Link accountability to action

Based on a review of relevant literature about statewide school accountability systems and findings from the case studies detailed below, this paper argues that the highest purpose of statewide school accountability systems should be to encourage specific action to improve school-level outcomes. With its focus on identifying and providing additional resources for the lowest performing schools,

ESSA presents a theory of action for statewide school accountability systems that attempts to connect accountability and school improvement (Domaleski et al., 2024). Given the importance of this connection, we place this idea at the top of our model as the ideal that statewide school accountability systems should strive toward.

Hierarchy of Purposes of Statewide School Accountability Systems

Level 1: Identify the purpose and vision for the education system.

Many preconditions and qualities must be present in a statewide school accountability system to create a connection between the system and school improvement practices. Foundationally, states must identify the purpose and vision for the larger education system that the accountability system will represent. SEAs often articulate the purpose and vision of the education system through high-level strategic plans. These plans are often complemented by theories of action and logic models that chart the path from high-level priorities and goals to more specific action.

Level 2: Design the system to best serve the purpose and vision.

Building on the vision for an education system, the SEA must then design the statewide accountability system in a way that best serves the goals of the strategic plan. System design includes two main categories of elements that must be considered in creating a sound statewide school accountability system.

- **System design** decisions identify the best approach to represent the metrics and data that are most relevant to understanding the performance of the education system. These decisions must balance several tensions to produce actionable outcomes. These tensions are covered further in the following section.

- **Technical design** decisions lay the framework for accurate data reporting on the metrics chosen for the accountability system. Indicators need to be designed in a way that encourages quality data collection, minimizes unintended consequences of including specific metrics in the system, and reduces data collection burdens on schools and districts as much as possible.

One prominent example of a common technical design decision comes in the weighting of student academic proficiency and growth measures, which is discussed in more detail in the following section.

Given the technical requirements of creating metrics, statewide accountability systems often cannot capture every element that is important in evaluating the performance of the education system. Rather, accountability systems must find a way to discern key elements that can identify schools and districts in need of additional support in a way that engenders support in the system (Olson, 2025). SEAs must also consider how these design decisions will impact the ways interested groups will interact with the system.

Level 3: Clearly communicate accountability information to a broad set of interest holders.

SEAs must also communicate data from the statewide school accountability system in a way that allows and encourages important groups to interact with the information. At the most basic level, ESSA requires that SEAs publish school report cards that present information about schools and districts. But because of the broad interest in education system performance across many groups, this information often needs to be communicated at many different levels (Munyan-Penney et al., 2024) beyond school report cards. Additionally, statewide school accountability systems communicate in both implicit and explicit ways. The inclusion and exclusion of certain metrics and how the SEA interprets the included metrics implicitly communicates certain values and messages about what is important within the education system. On the other side, the tools and methods used to communicate accountability outcomes (such as dashboards and processes for district- and school-level review of data) create the explicit ways that most interest groups interact with the accountability system.

Level 4: Create linkages between accountability information and action in order to improve outcomes.

Moving from communication to action within the education system requires SEAs to build on the more foundational levels described previously. Ideally, statewide school accountability systems can serve as a lever that drives and coordinates actions across numerous complex systems to make progress toward goals. Coordination can be thought of as occurring in two dimensions.

- **Horizontal coordination** across the SEA is required to align the actions of multiple systems, including (but not limited to) school improvement, resource allocation, and student support services.

- **Vertical coordination** between the SEA, districts, and schools is necessary to articulate the roles and responsibilities each level of the system plays and has in moving toward improved outcomes. Even in states that have a history of high levels of autonomy for LEAs, state-level coordination plays a key role in effective and efficient continuous improvement efforts.

The recent success of Mississippi in increasing student literacy rates illustrates the power of combining horizontal and vertical coordination to drive action. At the state level, the SEA aligned both efforts across agencies (horizontal coordination) to focus on providing coordinated statewide curricula and programming and supports for districts and schools with the highest need through the passage of the Literacy-Based Promotion Act (Breazeale, 2024). At the same time, the SEA invested heavily in building collaboration across all levels of the education system so that educators, administrators, and policymakers understood their role in creating change and had the necessary support to act. Mississippi's former Superintendent of Education described this process of vertical coordination, saying,

> Educators do not call these achievements a "miracle" because we know Mississippi's progress in education is the result of strong policies, the effective implementation of a comprehensive statewide strategy, and years of hard work from the state to the classroom level. (Excel*in*Ed, 2023)

While other states have adopted many of the tools Mississippi used to achieve statewide improvements, not all have had similar success, and this highlights the necessity of coordination and collaboration in addition to resources to drive change forward.

Tensions Within Accountability Systems

Based on our interviews and document review conducted in the case studies, we distilled four key tensions as states attempted to implement innovative accountability approaches. Within the context of these cases, addressing these tensions in meaningful ways often moved a statewide school accountability system closer to improving integration within the education ecosystem. On the other hand, a lack of clarity around these points can create additional challenges in an accountability system achieving its goals.

Tension 1: General Versus Targeted Use of the Accountability System

Even accounting for a variety of federal and state regulations that provide constraints, accountability systems have many flexibilities and can be designed in many ways while maintaining compliance with federal and state statutory requirements (Gong, 2024). In the cases we examined, states that set appropriately targeted goals were often more effective in communicating the goals, purposes, and limitations of the state's accountability system. Given the wide variety of priorities of accountability system users (state school improvement offices, LEAs, policymakers, etc.), creating a framework that establishes targeted uses can create common ground for the use of the system.

Finding the right scope for targeted uses of the accountability system can be tricky. If targeted uses are too high level and disconnected from on-the-ground contexts, many system users may not understand how the system attempts to coordinate actions. On the opposite end, designing a system in too specific a fashion can limit the impact of the system. For example, narrowly crafting your accountability system around highly technical school improvement data may provide districts and schools

with the detailed information needed to inform school improvement efforts, but it could discourage use of the accountability system by groups who typically take a less technical approach to the data (policymakers, parents, etc.). Building targets at the appropriate level can foster coordination across a wide variety of groups needed to further school improvement efforts.

Tension 2: Accessible Versus Complex Information

As seen in the case studies, SEAs have tended to add complexity to their statewide school accountability systems since ESSA was enacted a decade ago. On the one hand, this has tended to increase the amount of data available within accountability systems. On the other hand, it has also increased the complexity of the system, making it more difficult to understand.

Additionally, many states have designed accountability systems to provide summative school ratings such as school grades or star systems. These ratings often receive the most attention of any metric within the accountability system because they produce what appears to be a single, simple measurement of school quality. But while these ratings appear simple on the surface, they are often the result of a complicated series of decisions and calculations that distill a widely varying set of metrics into a single piece of information. While summative ratings often receive outsized attention among families and policymakers and in the media, these ratings provide little actionable information to district and school leaders. Finding the right balance between accessibility and complexity can help statewide accountability systems strike the right balance between the communication and technical needs for systemic school improvement.

Tension 3: Normative Versus Criterion-Based Approaches

Articulating and providing reasoning for specific approaches to the presentation of information is vital for the success of statewide school accountability systems; the choice between normative and criterion-based approaches is one of the most important decisions. Implementing a balanced and intentional process that

incorporates both approaches is necessary for accountability systems to incentivize action that improves outcomes.

Normative approaches compare the performance of students and groups of students to other students or groups but remain silent regarding the content knowledge students have. Utah's indicator that examines the performance of the bottom 25 percent of students at each school is a good example of a normative approach. This process is driven by a ranked ordering of students at each school as opposed to a measure of how much each student knows. Therefore, the performance against standard of a student at the 25th percentile at one school can vary significantly from that of a student at the same percentile in another school. This can be a good approach for continuous improvement because it can provide every school with a targeted group of students for whom to provide additional support, no matter the overall performance of the school. However, normative measures can be more complicated to communicate, particularly with school leaders and educators who focus their daily work on standards-driven knowledge growth in students.

Criterion-based approaches are the mirror image of normative approaches; they focus on student performance as compared to content standards rather than to the performance of other students. Proficiency measures are the most common form of criterion-based measures because they look at the number of students who achieve an agreed-upon standard (usually measured by a summative assessment) in a specific subject. Criterion-based metrics are usually less complicated and easier to communicate than normative metrics. However, criterion-based metrics often require a broader agreement on the articulation of the standards and essential skills students should master, which can be difficult to achieve considering the technical psychometric process that goes into developing those measurement tools (summative assessments are one good example of this). Overreliance on criterion-based metrics with a single cut score (e.g., proficiency rates) also can incentivize schools and districts to focus on students near the cut score to maximize accountability system benefits. This can have detrimental effects for students both above and below the cut score.

Tension 4: Statewide Requirements Versus Local Flexibilities

Every SEA must navigate the tension between creating statewide mandates and providing districts control to respond to differences in local contexts. Often referred to as "local control," SEAs must navigate the political environment within the state to implement the level of coordination necessary for creating fertile ground for improvement. They must also remain flexible enough to allow districts latitude to operate in the manner best suited for the districts' communities.

Strong requirements at the state level can encourage the creation of clear goals and alignment. This is useful in consistently important areas that require alignment across all districts. For example, in recent years SEAs have tended to increasingly tighten state-level requirements around literacy and numeracy curricula, instruction, and outcome metrics to coalesce focus on those fundamental skills. However, overuse of top-down priority setting can create a sense of detachment in users of accountability systems, which commonly results in a compliance-oriented relationship between users of the system and the system itself. Additionally, because state-level requirements must be implemented with fidelity across a wide variety of contexts, metrics can fail to capture the complexity of outcomes a school is trying to produce.

Creating a system with higher levels of local flexibility can generate more buy-in from communities and individual users of statewide school accountability systems. This increased investment can result in accountability systems that more accurately reflect the complexity and nuance of a community's goals for its education system. However, as complexity and nuance increase at the local level, outcomes become harder to compare across schools and districts. This comparability between schools and districts lies at the heart of federal accountability regulations in ESSA. Additionally, enacting state- and federal-level policies that ensure the quality of local accountability system implementation while preserving the flexibilities that produce community buy-in and innovation is often quite challenging to capture in state-level policy.

Linkages Between Hierarchy and Tensions

The hierarchy and tensions we have detailed operate in slightly different ways. The hierarchy is meant to highlight both the highest level at which a system (or portion of a system) operates and that the levels are cumulative. For example, if an accountability system is operating within the third level of the hierarchy (clearly communicating accountability information to different groups), it is still fulfilling the requirements of lower levels. In other words, the system is still operating from a clearly articulated vision and goals for the education system and has implemented a design that serves the vision and goals. Successfully implementing these lower levels is necessary for the system to operate at the higher level. For instance, communicating well with interested groups about accountability information requires both a clearly articulated vision for the system and a system that can connect the information back to the vision.

Additionally, SEAs may intentionally choose to return to lower levels of the hierarchy from time to time to reevaluate and strengthen foundational levels of the accountability system. A state engaging in strategic planning exercises provides a good example of an accountability system intentionally returning to a lower level. Periodic and intentional review of the foundational elements of an accountability system can ensure that the system operates in a way that best serves the purposes of the education system as changes inevitably occur over time (Crane, 2023).

Tensions interact with the hierarchy but operate differently. Each of the four tensions operate across the levels of the hierarchy, though the tensions look different based on the level at which the accountability system is operating. For example, the first tension (general versus targeted uses of the accountability system) can be used as framing for an SEA in the process of updating the vision and goals for the state as they go through a strategic planning process. If most students in the state are facing specific challenges, identifying more targeted goals for the accountability system may be in the best interest of the state. The Mississippi case mentioned above is a good example of a state pursuing literacy improvement in a more targeted way to focus attention, resources, and support on a specific issue.

As a state's accountability system moves up levels of the hierarchy, the general versus targeted tension is still present, but it manifests differently. Continuing with the example from above, as a state moves up the levels of the hierarchy to the fourth level (creating linkages between accountability information and action), the general versus targeted tension remains, but it does so in a different manner than at lower levels of the hierarchy. In Level 4, this tension may manifest itself in balancing resources devoted to specific targets, like literacy efforts, with more general supports for districts and schools. Given that SEAs work with finite resources, often targeted uses of resources take away from general supports. States must find the balance of targeting specific goals without diminishing resources to the point that the overall education ecosystem begins to suffer.

Specific tensions may also be more or less important depending on the level on which the system is operating. For instance, the third tension (normative versus criterion-based) might be only a background consideration as the state identifies its vision and goals, but it may become a central decision point as the state designs its accountability system and communicates information. Additionally, resolving tensions at lower levels can enable higher level functions. States that identify clear roles and actions to create coordination across tiers of the education system at lower levels of the hierarchy build the base for taking action at the top of the hierarchy. Returning to the Mississippi example, identifying clear roles for the SEA, LEAs, and schools as the state identified its vision and built systems allowed for tighter implementation of interventions to improve literacy outcomes.

Case Studies

To better understand the broader context of innovative school accountability approaches states have taken under ESSA, the project team identified three practices to examine within this research. The project team then identified four states that utilize these practices to develop the case studies described in this section. These practices and states are identified in the following list.

- Use of innovative accountability indicators

  - Ohio's use of an early literacy indicator within the state accountability system's School Quality and Student Success (SQSS) indicator

  - Utah's use of the growth of the bottom 25 percent of students as an indicator in the state's accountability system

- Use of dual accountability systems

  - Nebraska's use of a state-level accountability system, titled Accountability for a Quality Education System, Today and Tomorrow (AQuESTT), in addition to the accountability system required by ESSA

- Use of local flexibilities in accountability

  - Kentucky's development of community-driven accountability systems and attempts to align state policy with local flexibilities

These cases were not selected to provide a set of best practices for states to follow in revising their own accountability systems. Rather, these cases are presented as illustrations of how states have situated their accountability systems within the hierarchy and navigated the tensions presented.

Case Development, Data Collection, and Analysis

Using the research questions and guiding principles, the project team developed the structures for identifying the subjects of the case studies, collecting the data, and analyzing the gathered information.

Case Study Selection

Given the depth of information each case study needs to convey, one of the project team's most important tasks was selecting appropriate and meaningful instances of innovations. The project team developed a set of criteria to select case studies and scored all case study candidates according to the following processes.

Criteria for Case Study Selection

Using the research questions and guiding principles as a foundation, the project team identified five criteria to evaluate potential case studies.

- **Relevance:** The case study must directly connect to the research questions and provide an adequate depth of information to provide rich responses to each research question.

- **Innovation:** The case study should identify innovative approaches to accountability, defined as the practice not being widely adopted across the country and representing a new approach to accountability that goes beyond the explicit language of ESSA.

- **Diversity:** Taken as a whole, the case studies should represent a broad range of state contexts (size, geographic location, etc.) and perspectives.

- **Transferability:** Each case study should identify approaches that could be realistically taken on by a wide range of other states and organizations.

- **Data availability:** Each case study should present a variety of avenues for analysis, including artifacts that provide rich details and access to decision-makers who can speak to the full depth of the case.

Process for Selecting Case Studies

Members of the project team collaborated to identify potential case study candidates. In addition to the project team's collective knowledge of accountability systems across the country, the team also utilized resources such as the inventory of state accountability systems provided by the Education Commission of the States and additional literature about state accountability systems to identify as many possible candidates for the case studies as possible.

After collecting the case study candidates, the project team grouped case study candidates by topic. Each member of the project team independently rated each case study candidate based on four of the selection criteria: relevance, innovation, transferability, and data availability.[1] The project team summed the scores for each of the four categories to identify a tier of case studies that scored the highest for

---

[1] Diversity was omitted as a rating at this point because this criterion required looking at the group of case studies as a whole. Providing a rating on an individual case level would have been counterproductive at this stage of the process.

each topic. Then, utilizing the total score and considering the diversity of the highest scoring cases, the project team selected the final set of case studies.

Methodology and Interview Protocol

To develop the case studies presented in this report, the project team employed a qualitative research approach, combining document review with semistructured interviews to generate in-depth insights into innovative accountability practices across four states. This mixed-method approach allowed for a nuanced understanding of both the technical and contextual factors that shape accountability reform efforts.

Data Collection and Interviews

Our data collection process consisted of two primary components:

1. **Document and artifact review:** For each case study, we gathered and analyzed key documents—including state accountability plans, legislative texts, technical manuals, implementation guidance, and publicly available data reports—to build foundational knowledge and contextualize each innovation.

2. **Semistructured interviews:** We conducted 11 interviews with 17 participants across the four states studied. These included

   - two interviews with Kentucky interested holders, including district leaders from Shelby County and Fleming County;
   - five interviews in Nebraska, involving both state officials and district administrators from urban and rural contexts;
   - two interviews in Ohio, primarily with leaders in the Office of Accountability and Curriculum; and
   - two interviews in Utah, including state assessment and accountability leaders and research staff.

Interviews were guided by a semistructured protocol designed to elicit detailed narratives aligned with the project's research questions and guiding principles. The

protocol included anchoring questions with opportunities for follow-up based on participants' roles and insights. The full interview protocol can be found in Appendix B.

Analytical Approach

Interview transcripts and artifact notes were analyzed thematically, with findings organized around the four guiding research questions. Special attention was given to triangulating themes across data sources and highlighting both the implementation processes and the policy implications of each innovation.

Limitations and Potential Bias

While this qualitative approach enabled a rich, practice-informed view of accountability system innovation, it carries inherent limitations.

- **Selection bias:** Participants were selected based on their involvement in the identified innovations. As such, perspectives may lean toward those more supportive or directly engaged in these reforms.

- **Scope of states:** The four states profiled were selected to highlight different types of innovation and represent geographic and contextual diversity. However, findings are not intended to be generalizable to all states.

- **Lack of counterfactual perspectives:** This study does not include interviews with critics or interested holders who may have had opposing views on the innovations presented. As a result, the perspectives reflected are primarily from proponents or implementers.

- **Absence of longitudinal impact data:** This study emphasizes perceived and early-stage impacts as described by practitioners and supported by available documentation. Long-term outcomes will require further evaluation.

Despite these limitations, the consistency of themes across states and interest holders suggests that the findings offer valid insights into the potential and challenges of state-level accountability reform.

Centering Early Literacy in Accountability: Ohio's Innovation and Implementation Experience

In recent years, early literacy has played a prominent role in state-level education reform, driven by a growing understanding of its critical role in long-term academic success. However, few states have successfully integrated early literacy into formal accountability frameworks, largely due to the technical and developmental complexity of measuring young children's reading progress.

Ohio's accountability system includes a notable innovation aimed at addressing long-standing disparities in foundational reading achievement: the early literacy indicator. This measure tracks how well schools are supporting students in grades K–3 in developing essential reading skills, including their progress toward reading at grade level by the end of grade 3. The early literacy indicator was introduced in response to growing concern, at both state and national levels, that students who do not achieve reading proficiency early are more likely to struggle academically in later grades.

Recognizing that early reading skills are critical predictors of long-term educational success, Ohio developed the early literacy indicator as a measure of foundational learning. By focusing attention on students who need additional support in the earliest grades, the early literacy indicator is intended to close achievement gaps before they widen. It reflects the state's broader goal of ensuring that every student, regardless of background, has access to strong early learning opportunities and interventions.

The Early Literacy Indicator tracks improvement in three areas: kindergarten readiness, grade 1 and grade 2 literacy improvement, and grade 3 reading proficiency. This comprehensive approach was shaped through input from educators, district leaders, and state policymakers committed to building a system that both identifies need early and supports responsive action. By building such a system, Ohio aims to integrate accountability and support into the foundational years of student learning.

Context and Origin of the Early Literacy Indicator

Ohio's Early Literacy Indicator emerged from a broader state commitment to improving foundational reading skills and closing equity gaps in K–3 learning. One district administrator emphasized, "We know that early literacy doesn't start in third grade. It's progress then from kindergarten to first, and from first to second, and second to third." The accountability component formalizes this commitment by tracking key metrics across three domains.

- proficiency on the grade 3 English Language Arts (ELA) test

- promotion to grade 4

- improvement among students previously off track

The Early Literacy Indicator is uniquely situated within a larger constellation of early literacy reforms, including the adoption of the science of reading (SOR) framework, a required dyslexia screening law, and mandated professional development (PD) for teachers using state-approved literacy curricula.

Leaders at the Ohio Department of Education and Workforce, particularly in the Office of Accountability, worked closely with school districts and educators to design the Early Literacy Indicator, aligning it with existing literacy screening requirements and implementation plans. One leader noted, "We are working hard on building that coherence, which will be the accelerant toward improved outcomes for kids. … Accountability can't happen if we don't have coherence."

Connections to the Hierarchy and Tensions

Ohio's Early Literacy Indicator is a good example of a lower level change to an accountability system impacting higher level functions. While the change itself would be classified as a change to the design of Ohio's accountability system (Level 2) to better connect to the purpose and vision of the education system (Level 1), the implementation of this innovation had effects on both communication and connections to action.

In implementing the Early Literacy Indicator, Ohio had to navigate two main tensions. First, Ohio identified that early literacy improvements warranted action that was more targeted within the accountability system, as opposed to a more general approach that may not have met the state's goals. Second, by adding an additional indicator, the accountability system became more complex. While the added complexity could make the system more difficult to navigate, officials in Ohio determined that the deeper information would move the state closer to achieving its larger goals.

Impact of the Early Literacy Indicator

According to school and district leaders interviewed for this case study, the inclusion of early literacy in accountability has created new urgency and clarity around reading instruction and intervention. Educators now view literacy screeners not as isolated compliance exercises but as tools that are foundational to school improvement strategies. As a district administrator opined,

> I think that [the Early Literacy Indicator] coupled with the science of reading, and state dyslexia law, where you have to implement a state-approved assessment for that on-track, off-track piece when accompanied by [data literacy] professional development—I think that has been really powerful.

One district leader described this as a "trifecta" effect—early literacy indicators working in tandem with the SOR and dyslexia mandates to elevate the visibility and intentionality of literacy instruction across the K–3 continuum.

Early outcomes include

- a jump in early literacy proficiency from 76.6 percent in 2021/22 to 93.2 percent in 2023/24 in one district;
- a stronger data culture in elementary schools, with teacher teams using screener data and state report card metrics to guide interventions; and
- the creation of districtwide data protocols and early warning systems, helping educators act on diagnostic information in real time.

Educators emphasized that the early literacy indicator's impact is amplified by accompanying supports, particularly targeted PD and flexible guidance from the state.

Implementation Successes and Challenges

As Ohio integrated the early literacy indicator into its accountability system, the perspectives of educators and district leaders highlighted both the promising shifts in early grade instructional focus and the practical hurdles involved in measuring and improving foundational reading outcomes at scale.

Successes

- **Capacity building and data literacy:** State investments in PD, including online modules and in-person coaching, helped educators learn how to interpret screener and assessment data. This was crucial for ensuring that accountability measures translated into instructional change.

- **Alignment across systems:** Teachers and administrators noted how the alignment of the early literacy indicator with SOR and dyslexia initiatives made the system feel coherent. As one leader put it, "This wasn't just another requirement … it was connected to everything else we were doing."

- **Community communication:** The public-facing nature of the early literacy indicator has strengthened family engagement. Educators reported greater fluency in using the report card to explain school priorities to parents.

Challenges

- **Complexity of screening tools:** With multiple state-approved vendors, some districts struggled to align local tools with state reporting expectations. Differences in cut scores and timing added confusion in interpreting progress.

- **Data integration:** Leaders stressed the importance of having a centralized data system and highlighted challenges arising from the lack of infrastructure.

- **Sustainability concerns:** Some educators worried that proposed budget cuts might limit ongoing access to PD and instructional materials, threatening the durability of current gains.

Policy Implications

Ohio's experience offers concrete insights for states considering the inclusion of early literacy in accountability.

- **Treat literacy accountability as part of a broader ecosystem.** Literacy indicators are most effective when paired with aligned mandates (e.g., dyslexia screening), funding, and instructional guidance.

- **Fund and require literacy-focused PD.** Ohio's success was catalyzed by a state policy that not only required PD but also paid teachers to complete it. Other states should follow suit by embedding funded PD requirements into literacy accountability rollouts.

- **Prioritize data literacy and infrastructure.** Accountability reforms must be accompanied by tools and training. Ohio districts reported the importance of data warehouses, visualization platforms, and designated literacy/data coaches.

- **Balance rigor with flexibility.** Ohio's approach combined clear expectations with choice, allowing districts to select screeners from a vetted list and adapt state guidance to their local needs.

- **Leverage public reporting for transparency and engagement.** The presence of the early literacy indicator on public report cards helped educators communicate with families and advocate for resources.

- **Evaluate for impact.** As the early literacy indicator matures, ongoing research should examine how early literacy outcomes change across the state, with a particular focus on variance across subgroups.

Reframing Academic Growth: Utah's Use of the Lowest Quartile Indicator

Across the country, states are exploring how accountability systems can better serve the goal of producing improved outcomes for the lowest performing students. While compliance and comparability remain essential, many states, like Utah, are redefining accountability to emphasize support, growth, and reduced achievement gaps. Utah's accountability model incorporates a distinctive innovation: a growth indicator focused on the lowest performing 25 percent of students in each school.

Utah's innovation emerged in the wake of national and local recognition that status-based accountability systems, focused primarily on proficiency, often fail to recognize the progress of students who are furthest from meeting grade-level expectations. In particular, educators and policymakers in Utah expressed concern that the state's prior models were not adequately capturing the growth of students performing in the lowest quartile of academic achievement. Interest holders sought a more meaningful way for schools, no matter their overall performance level, to focus on the needs of the lowest performing students in the school to foster a more robust continuous improvement process.

As part of its broader Next Generation School Accountability initiative, Utah sought to elevate equity not just as a guiding principle but also as a measurable outcome. State statute (Utah Code 53E-5-205) required the inclusion of a measure of equitable educational opportunity, creating a mandate and opportunity to develop the lowest quartile (LQ) growth indicator. The goal was to move beyond static, one-size-fits-all indicators and instead shine a spotlight on how well schools were serving their most vulnerable learners. By doing so, the state aimed to focus school and district attention on students with the greatest need for support, encourage inclusive instructional practices, and drive targeted improvement strategies statewide.

Context and Origins of the LQ Indicator

The development of Utah's LQ indicator was shaped by a deliberate effort to translate the state's equity commitments into actionable, data-informed policy.

Rather than adding a new metric in isolation, state leaders built on prior conversations within the Assessment and Accountability Policy Advisory Committee and the broader ecosystem of interest holders. These discussions consistently raised the need for transparent, actionable ways to assess school impact on students who face the steepest barriers to academic success.

The LQ indicator was designed with these practical concerns in mind. Education leaders wanted a measure that would highlight the progress of students in the lowest quartile without relying solely on subgroup performance categories. Using student growth percentiles (SGPs), the LQ metric captures the share of students in a school's lowest quartile who achieved typical or better growth from one year to the next. This allows schools to be recognized for helping all students make progress, not just those near proficiency thresholds.

This approach reflects a broader commitment to continuous improvement. The LQ indicator was piloted and refined through Utah's Next Generation School Accountability vision, which prioritizes clarity, coherence, and interest holders' ownership in system design. State education staff, assessment directors, and school leaders all played roles in adapting the indicator to local realities. One state-level accountability administrator pointed out that "there was a very large group of stakeholders … principals, teachers, politicians, psychometricians from the Center for Assessment. … They outlined principles of what do we want to be held accountable for." Their feedback helped ensure that the metric would not only meet statutory requirements but also support meaningful instructional decision-making in diverse educational settings.

Connections to the Hierarchy and Tensions

Like Ohio's early literacy indicator, the primary mechanism of Utah's LQ indicator serves to align the vision and goals of the state's education system with the design of the accountability system. The LQ indicator also attempted to serve as a bridge to higher level actions by spotlighting the performance of students that needed the

most support. Further actions by the state leveraged the LQ indicator as a communication tool that highlighted areas for further actions by districts and schools.

In implementing the LQ indicator, officials in Utah had to navigate two main tensions. First, the indicator includes ideas that are more complicated than many other indicators within the system. For example, the idea that the level of the bottom 25 percent of students as a group will vary from school to school may not be an intuitive idea to all audiences. Officials in Utah made an intentional choice to utilize a normative approach within this indicator, comparing students to other students instead of to a set performance level. This approach allows for the identification of the lowest performing students at each school and provides useful information for continuous improvement at all schools in a way that may not have been possible using only criterion-based information.

Impact of the LQ Indicator

Although Utah has not published a formal impact evaluation of the LQ indicator, evidence from practitioner interviews, Utah State Board of Education (USBE) documentation, and district practices suggests that it has shifted how schools talk about data and student growth. Several assessment directors and school leaders noted that the LQ indicator helps illuminate gaps that might otherwise be hidden by aggregate data.

Educators report that the LQ indicator prompted changes in how they structure data conversations. Rather than focusing solely on average growth or proficiency, teams now regularly disaggregate results and ask, "How are our struggling students progressing?" For some schools, the LQ indicator became a rallying point for tiered intervention systems and targeted instructional strategies. Another state-level accountability administrator emphasized,

> When I was a department chair … that was one of the big things that we
> focused on to start implementing schoolwide procedures on how to reach

the students that weren't getting it the first time … it definitely helped for my little middle school out in Riverton.

Importantly, interest holders including principals and union leaders have emphasized the motivational power of growth-based indicators over status-based indicators. Since the LQ metric highlights improvement rather than raw scores, it fosters a culture of possibility and encourages schools to focus on supporting growth for every student.

Implementation Successes and Challenges

As Utah's LQ indicator moved from design to implementation, the experiences of educators and system leaders revealed both the promise of this measure and the practical complexities involved in embedding it within school improvement efforts.

Successes

- **Clarity of purpose:** Educators understood the LQ indicator's intent and appreciated its alignment with equity goals. The USBE's technical guidance clarified how the indicator works and why it matters.

- **Cultural shift:** The metric encouraged a move from deficit-based thinking to growth-oriented practice, particularly in schools previously labeled as "failing."

- **Capacity building:** Professional learning offered by the USBE helped districts integrate the LQ measure into their continuous improvement planning.

Challenges

- **Understanding and communication:** Not all interest holders initially grasped how the LQ group was calculated or how growth targets were set. Some parents and educators viewed the term "lowest 25 percent" as stigmatizing.

- **Data complexity:** Reliance on SGPs created challenges in interpretation, especially in schools where year-over-year group sizes vary.

- **Scale of use:** While district leaders often use the LQ indicator for planning and support, it is underutilized by community interest holders and remains absent from some public reporting narratives.

Policy Implications

Utah's use of the LQ indicator provides a model for how growth-focused measures can elevate accountability systems. Several policy lessons emerge.

- **Elevate growth to emphasize continuous improvement.** State-level policy can emphasize the use of growth-based indicators that balance proficiency and academic growth. This approach could incentivize schools to focus on a broader set of students and create a more robust culture of continuous improvement.

- **Balance simplicity and rigor.** Technical metrics like SGPs must be paired with strong communication with interest holders.

- **Invest in process, not just outcomes.** Utah's multiyear process of advisory input, technical refinement, and educator engagement demonstrates that innovation is as much about how a policy is implemented as it is about what it measures.

More than One Way to Get There: Nebraska's Use of Dual Accountability Systems

In a handful of states, SEAs have developed statewide accountability systems separate from the federally mandated system. These state-specific accountability systems operate in parallel with the federal system with varying levels of coordination, depending on the state. Nebraska provides an example of a state running two accountability systems concurrently with its AQuESTT operating alongside the federal school accountability system within the state.

States have a variety of motivations for running parallel statewide school accountability systems. Some states had existing accountability systems prior to the enactment of ESSA and chose to maintain that system in addition to fulfilling federal

requirements. Other states enacted state-level systems to provide additional information and flexibility not found within the federal regulations. Still other states were required to build parallel systems by statutes passed by state legislatures. Regardless of what led to the development of these parallel systems, several states have pursued this avenue to improve school accountability and improvement practices. Nebraska's experience implementing the AQuESTT system provides deeper insight into the motivations, impact, and challenges of building and executing parallel systems at the state level.

Context and Origins of the AQuESTT System

The AQuESTT system was first implemented in 2015, prior to the passage of ESSA. The system was shaped by the passage of the Quality Education Act in Nebraska, which set the minimum standards for accountability and provided the state board with the authority to develop its system. Initially, AQuESTT operated on a 3-year cycle, establishing baseline years before reclassifying schools every three years. However, ESSA's requirement for annual school identification necessitated AQuESTT shifting to annual classifications.

In addition to being implemented prior to the passage of ESSA, AQuESTT also offers marked contrasts to ESSA in terms of how it approaches school accountability. While both the AQuESTT system and ESSA requirements involve a series of indicators to measure school performance, AQuESTT grounds these indicators within a framework of six tenets (Nebraska Department of Education [NDE], n.d.). These tenets provide an explicit model for school improvement that ESSA regulations do not articulate. This ability to define and more closely connect a model of school improvement to the accountability system is one potential strength of state-level accountability systems.

In addition to better connections to an actionable school improvement framework, AQuESTT actively avoids the practice of ranking schools to identify schools in need of additional support. As Nebraska's Director of Accountability emphasized,

> Because the nature of ESSA is essentially ranking and knowing that we got such pushback against ranking ... I broached that with one of our advisory committees and the response was overwhelmingly no, because they felt like if a system is supposed to be used for ranking, it could be used for ranking, and they didn't want that.

Instead, schools receive scores for each indicator that are compiled into a final classification that determines the levels of support and intervention the school receives. In other words, while ESSA requires a minimum percentage of schools to be identified for additional support, there is no such requirement in AQuESTT.

AQuESTT has undergone numerous revisions in the decade it has been in existence, not only to allow the state to remain in compliance with federal ESSA requirements but also to strengthen the connections with school improvement practices. Currently, a group of district-level administrators is leading another round of adjustments to the system, including emphasizing the use of criterion-based metrics with the goal of drawing direct connections to school improvement efforts.

Connections to the Hierarchy and Tensions

Nebraska's utilization of a state-level accountability system largely focuses on the top two levels of the hierarchy presented in this paper. By presenting AQuESTT as explicitly connected to a framework for school improvement, the state is attempting to establish common terminology and understanding about the connection between the two areas, which can improve the clarity of communication across many audiences. Additionally, this positioning of AQuESTT also attempts to drive action for school improvement.

This case also draws out the tension around normative versus criterion-based information. Schools and districts often operate more on a criterion-based information level–for example, how much does a student know compared to a standard for someone at that grade level. However, many accountability systems, especially under ESSA, use normative information that is tied to how other students

or schools perform rather than using content standards. Bridging this gap is a key tension that must be addressed within any accountability system that hopes to build a connection to school improvement.

Impact of the AQuESTT System

Situating an accountability system intentionally in a framework geared toward school improvement has created an explicit connection between the two areas. Several district administrators mentioned that the framework behind the AQuESTT system has established shared terminology that they use to communicate with a wide variety of audiences. As one district administrator said, "I always take the time to share AQuESTT results with our entire staff, our school board, and the local media. The classification we receive allows for a good starting point in the conversation about school improvement."

While most district administrators we talked to agree that AQuESTT provides a good starting point for the school improvement conversation, deeper connections to improved outcomes have been tougher to achieve. While current revisions to the AQuESTT system will attempt to make information more understandable and actionable for districts and schools, previous versions of the system struggled in this regard. According to one district administrator, "Taking a normative approach to information is a hard thing for schools and districts to get behind because it's difficult to make the goals and metrics concrete. We think in ways much more aligned with a criterion-based approach." Additionally, the complexities of implementing both a state- and federal-level system can require additional time in releasing the data. One district administrator highlighted the timing mismatches between accountability and school improvement systems: "The public release of accountability information has typically taken place just before Thanksgiving. Although we get some of that data earlier, as a district leader you're always working half to three-quarters of a year behind."

Implementation Successes and Challenges

Given the added complexity of administering two separate statewide accountability systems, it comes as no surprise that the NDE encountered both successes and challenges in implementing AQuESTT.

Successes

- **Inclusive development and revisions processes:** The NDE has been intentional about including a wide variety of educators and interest groups in developing and revising AQuESTT. While many involved in this process agree that the system is not perfect, the process has built trust and buy-in.

- **Common language regarding school improvement:** AQuESTT introduced shared language and definitions regarding school performance and improvement, which has allowed for clearer goal-setting processes across the state. The relative longevity and consistency of the system has allowed for greater saturation of key concepts.

Challenges

- **Complexity of communication:** While the NDE has attempted to bridge differences between AQuESTT and the federal accountability system, messaging about how the two systems interact persist. Managing communication and processes around school identification and the resulting next steps for districts and schools has been particularly challenging.

- **Misalignment in data releases and school improvement planning:** Many states struggle to release data from statewide accountability systems in time to inform school improvement planning. The added complexity of running two systems further stresses this misalignment in Nebraska because formal releases of information do not occur until well after the school year starts.

Policy Implications

Nebraska's implementation of dual accountability systems offers several policy implications for other states already implementing dual systems or states considering this approach.

- **Detailing connections between accountability and school improvement can improve communication.** One of the greatest strengths of implementing dual accountability systems is the potential to draw clearer connections between accountability information and key school improvement concepts. This can enhance communication and improve utilization of accountability information.

- **Data and communication must be actionable.** Detail and nuance of an accountability system must be balanced with delivery of this information in a way that encourages action. Timing of communication and data release is a key component determining how well districts and schools can drive action from the information.

- **Thoughtful engagement builds trust.** As seen in both this case and the Utah example, accountability systems can be invaluable tools in engaging with important interest holders to build trust.

Balancing Statewide Standards and Local Innovation in Kentucky

Across the United States, state accountability systems have long served as mechanisms to measure school performance, identify areas for improvement, and ensure that students are receiving equitable educational opportunities. While originally rooted in federal compliance requirements, such as those outlined in ESSA, accountability systems increasingly aim to foster continuous improvement, inform instruction, and reflect local priorities. But the transition from compliance-driven to improvement-oriented accountability presents inherent tensions between standardization and flexibility, federal mandates and local autonomy, and academic outcomes and holistic student development.

Kentucky's integration of local control and public data transparency offers a compelling example of how lower level changes in accountability system design can drive broader shifts in practice and culture. While these changes are rooted in design-level innovations, such as the inclusion of locally developed indicators and district-led data dashboards, they have had cascading effects on how data are communicated and used for action across the state.

Through its United We Learn initiative, Kentucky launched the Local Laboratories of Learning (L3) to support community-driven, student-centered accountability innovation. Districts like Fleming County and Shelby County were among the first to pilot this work, developing locally tailored accountability models that include performance-based portfolios, interdisciplinary assessments, and student exhibitions. These models are grounded in each district's Portrait of a Learner and emphasize durable skills like communication, problem-solving, and self-regulation—skills identified by local communities as essential for college and career readiness.

To ensure data transparency and build trust, Kentucky now requires every district to publicly post student achievement outcomes, and many districts have developed public-facing data dashboards to make this information accessible and actionable. In Shelby County, for example, the district created school-based dashboards that feed into a broader district dashboard; these dashboards were shaped directly by input from over 30 interested party groups, including students, parents, and business leaders. One district administrator shared, "We met with 30 groups—students, community members, businesses. They want to see growth and achievement. They don't care about the rest of the metrics or the ed-specific language." These dashboards prioritize information that communities care about most—growth and achievement—and avoid technical jargon that can obscure understanding.

At the school level, districts like Fleming County have also implemented individual student portfolios for all students in grades 3–12, showcasing academic and nonacademic growth through personalized artifacts. These portfolios are now used for promotion and graduation decisions and are scored using validated, codeveloped rubrics. Each school in Fleming is now piloting its own local accountability model.

"We've not given any restrictions to them about what that model looks like," noted a district superintendent. "They need to engage their communities, their parents, their students, and their teachers." These school-level models include dashboards, community stories, and public-facing performance data.

Together, these innovations reflect a strategic shift in Kentucky's accountability system, one that balances the state's equity goals and federal requirements with the autonomy and insight of local communities. The result is a more nuanced and transparent system that not only informs practice but also strengthens the connection between schools, students, and the communities they serve.

Context and Origins of Innovation

Kentucky's accountability redesign emerged from a desire to move beyond standardized, compliance-oriented assessments and respond to local educational values and needs. The impetus for innovation was grounded in the feedback of interest holders, equity goals, and a recognition that traditional accountability structures often failed to reflect the full scope of student learning. The Kentucky Department of Education (KDE), in collaboration with the United We Learn Council, launched a series of design prototypes shaped by educators, families, students, and community leaders. Educators from across the state played a central role in this transformation, contributing insights based on their practical experiences with the shortcomings of existing systems.

Structures such as L3, town halls, and educator advisory groups were used to ensure that those implementing the innovations had a voice in their design. Their feedback helped clarify what was feasible given varying district capacities and surfaced scalable innovations that met local needs. These ranged from district-specific portfolios and student exhibitions to full-scale districtwide performance dashboards and competency-based assessments.

Connections to the Hierarchy and Tensions

Kentucky's community-driven accountability redesign is a strong example of a foundational, vision-setting effort that has begun to influence higher level functions within the system. While the work would be classified as a Level 1 change (establishing the purpose and vision for the education system), the state's emphasis on codesigning the Portrait of a Learner and engaging districts through structures like L3 has created the conditions for meaningful communication (Level 3) and early connections to local improvement actions (Level 4). By building trust and ownership through inclusive engagement, Kentucky has laid a durable foundation for more coherent and aligned accountability practices to emerge over time.

In developing this approach, Kentucky had to navigate the tension between statewide requirements and local flexibilities. Rather than enforcing uniform mandates, the state created space for districts to pilot customized performance measures aligned with the shared vision, such as exhibitions, student-led conferences, and district dashboards. This flexibility helped foster community relevance and professional buy-in but also surfaced challenges around comparability, data coherence, and sustainability. Without codification in state policy, Kentucky's locally driven model may face difficulties maintaining momentum and consistency as leadership or priorities shift. Nonetheless, the Kentucky case illustrates how vision-driven design can build the conditions necessary for innovation, and it highlights the importance of policy infrastructure to support long-term implementation.

Impact of Innovative Accountability Practices

While the full impact of Kentucky's redesigned system is still developing, early signals demonstrate how implementation processes have influenced both practice and culture within schools and districts. Districts such as Shelby County and Fleming County have documented shifts in instructional practice, data use, and educator collaboration. "Now the conversation with teachers isn't about a test score. It's about readiness—can this kid go to the next grade and succeed?" one Fleming County

administrator noted. "And it's not necessarily based on a test score. We now look at multiple measures of success." These changes are attributable not only to structural innovations but also to the inclusive, educator-centered approach the KDE used to support implementation.

For instance, Shelby County Public Schools staff developed robust internal structures to support data interpretation and continuous improvement. Educators there emphasized the importance of professional trust and autonomy in tailoring assessments to local goals. Similarly, teachers in Fleming County led efforts to develop the Measures of Quality framework, cocreating competency rubrics and guiding students through performance-based promotion requirements. Staff reported that this work deepened instructional relevance and provided clearer signals of student readiness.

Importantly, these innovations vary in scope and scale. While Shelby County implemented districtwide Profile of a Graduate exhibitions, other districts participating in early pilot phases may have introduced only one or two capstone projects or student-led conferences per school. This range of scope allows for customization depending on district capacity, culture, and readiness for change.

Implementation Successes and Challenges

Kentucky's approach centers the experience of educators and district staff in every phase of implementation, and their insights have shaped both successes and challenges.

Successes

- **Ownership through codesign:** Teachers and administrators led much of the design work, increasing buy-in and sustainability. The process created a culture of shared accountability and professional collaboration.

- **Process-oriented structures:** Triannual feedback reviews, supported by the KDE and peer districts, helped identify implementation strengths and gaps without rushing to premature outcome evaluations.

- **Differentiated implementation:** Districts scaled innovations based on capacity, from limited pilots to full frameworks, allowing for manageable change.

Challenges

- **Data management and use:** Many staff expressed concerns about the complexity of managing multiple data sources and ensuring consistent interpretation.

- **Balancing innovation and mandates:** Educators noted tensions between new local measures and traditional state assessments, which still heavily influence perceptions of success.

- **Capacity gaps:** Smaller or rural districts sometimes lacked the staff or resources to design new accountability measures, even with state support.

Policy Implications

Kentucky's experience provides rich lessons for federal and state policy rooted in the lived experiences of educators and district leaders.

- **Center educator voice in policy design.** Kentucky's success was made possible by designing an engagement process that allowed educators to provide valuable insights that informed system design.

- **Emphasize process over premature outcomes.** Federal and state guidelines should promote phased implementation timelines that allow innovations to mature before judging their impact based on student test scores.

- **Support scalable entry points.** States should offer a range of implementation options so that districts with varied capacity and appetite for change can engage meaningfully. Small pilots can lead to larger reforms.

- **Invest in capacity building and peer networks.** Kentucky's professional learning networks and peer review structures have been key to sustaining and scaling local innovations.

- **Foster alignment without uniformity.** Rather than enforce one-size-fits-all accountability models, policy should promote alignment around shared goals like equity and student agency while allowing districts flexibility in how they measure success.

Conclusion

As illustrated in the four studies presented in this paper, states have taken a variety of approaches in refining their statewide school accountability systems in order to better drive improvements in school and student outcomes. Though these methods have varied widely, the successes of these changes share a few common themes:

- Successful refinements were those made intentionally to align changes with the larger vision of the education system and design the accountability system thoughtfully to address the goals of that vision.

- Strong and coordinated communication with educators and other interest holders was a vital part of the success of the refinements states implemented within their accountability systems.

- Although states saw varied results in connecting accountability system refinements to school improvement efforts, each state articulated a strong theoretical model regarding the connection.

Although the context of the accountability systems varied across the states, these themes were present in each case, demonstrating the potential of the hierarchy presented in this work.

Similarly, each of the states grappled in some way with one of the four main tensions presented in this work. In some cases, states found a method to address the tension that led to greater coordination between the accountability system and efforts to improve student outcomes. In other cases, states were unable to find an effective solution to the tension, which resulted in continued disconnects between accountability

information and action. However, all states recognized the importance of addressing these tensions to create more effective statewide school accountability systems.

Although this work has identified promising commonalities both in terms of approaches to accountability system refinements and key tensions that can create challenges across these four states, additional work will be necessary to understand whether these themes and tensions apply to states more broadly. In addition, more study will be necessary to craft specific and actionable policy recommendations for accountability based on this model, particularly in accounting for the wide variety of state contexts across the country and the changing landscape of federal education policy. However, this model could provide a useful framework for deepening our understanding of statewide school accountability systems and the broader connection to effective school improvement efforts.

References

Breazeale, G. (2024). *Literacy series: Pre-pandemic and post-pandemic literacy achievement in Mississippi*. Mississippi First. https://www.mississippifirst.org/literacy-series-part-1/#%3A~%3Atext%3DBetween%202013%20and%202019%2C%20fourth%2Cappro aches%20to%20elementary%20school%20literacy

Burnette, D., II. (2018, June 12). When states and the feds take different paths on school accountability. *Education Week*. https://www.edweek.org/policy-politics/when-states-and-the-feds-take-different-paths-on-school-accountability/2018/06

Carlson le Floch, K., & Atchison, D. (2025, July 23). Is ESSA accountability playing out as intended? *American Institutes for Research*. https://www.air.org/resource/blog-post/essa-accountability-playing-out-intended

Crane, E. (2023, November 1). Reviewing state accountability systems with a beginner's mind. *WestEd*. https://www.wested.org/blog/reviewing-state-accountability-systems-with-a-beginners-mind/

Domaleski, C., Pinsonneault, L., & McCauley, C. (2024, May 1). Strengthening ties between accountability and school improvement: Five principles to guide practice. *The National Center for the Improvement of Educational Assessment*. https://www.nciea.org/blog/strengthening-ties-between-school-accountability-and-improvement/

EdTrust. (2023). *Joint comment regarding the Innovative Assessment Demonstration Authority*. https://edtrust.org/press-room/joint-comment-regarding-the-innovative-assessment-demonstration-authority/

Excel*in*Ed. (2023, August 11). Four reasons why Mississippi's reading gains are neither myth nor miracle. https://excelined.org/2023/08/11/four-reasons-why-mississippis-reading-gains-are-neither-myth-nor-miracle/

Gong, B. (2024, March 13). You can get the accountability system you want (Part 2): Beyond ESSA: Three state and district examples. *The National Center for the Improvement of Educational Assessment*. https://www.nciea.org/blog/get-the-accountability-system-you-want-part2/

Munyon-Penney, N., Jones, A., & Levitan, S. (2024). *Reassessing ESSA implementation: An equity analysis of school accountability systems*. EdTrust. https://edtrust.org/rti/reassessing-essa-implementation-an-equity-analysis/

Nebraska Department of Education. (n.d.). *AQuESTT: Accountability for a Quality Education System, Today and Tomorrow* [Infographic]. https://aquestt.com/wp-content/uploads/2020/07/2020AQuESTTFrameworkandTenets.pdf

Olson, L. (2025). *Quality check: The new, best way to measure school performance*. FutureEd. https://www.future-ed.org/quality-check-the-new-best-way-to-measure-school-performance/

Pinsonneault, L. (2023, March 15). Reframing accountability: Build closer connections between school identification and school improvement. *The National Center for the Improvement of Educational Assessment*. https://www.nciea.org/blog/reframing-accountability-build-closer-connections-between-school-identification-and-school-improvement/

Appendix A: Interviews

Date	State and district	Participants
1/17/2025	Ohio Department of Education	Rachel Wakefield, Interim Director of Accountability
2/11/2025	Utah Department of Education	Brittney Broadhead, School Accountability Specialist Cyd McCarthy, Director of Assessment and Accountability Megan Tippetts, Research Consultant III
2/12/2025	Nebraska Department of Education	Derek Ippensen, Director of Accountability
3/17/2025	Utah Department of Education	Aaron Brough, Director of Data and Statistics
3/18/2025	Nebraska, Hampton Public Schools	Holly Herzberg, Superintendent
3/20/2025	Nebraska, Millard Public Schools	Darin Kelberlau, Executive Director of Assessment
3/20/2025	Nebraska Department of Education	Lane Carr, Administrator, Office of Policy and Strategic Initiatives
3/25/2025	Nebraska, Scottsbluff Public Schools	Andrew Dick, Superintendent
3/26/2025	Kentucky, Shelby County Public Schools	Adam Hicks, Assistant Superintendent of Curriculum, Instruction, and Assessment Daniel Pfaff, Digital Learning Coordinator
3/27/2025	Ohio Department of Education	Katie Nowak, Director of Curriculum and Gifted

4/8/2025	Kentucky, Fleming County Public Schools	Brian Creasman, Superintendent Michelle Hunt, Chief Academic Officer

Appendix B: Interview Protocol

Gates Case Study Interview Protocol
Winter 2025

Interviewee Name, Agency:

Interviewer:

Date:

Introduction and Purpose

Hello, I'm _____ from WestEd [a recognized independent not-for-profit research and development agency]. WestEd is partnering with educational agencies to investigate innovative accountability practices within the framework established by the Every Student Succeeds Act (ESSA). This research aims to understand the experiences of State Educational Agencies (SEAs) as they implement flexibilities in accountability systems and to derive generalizable recommendations for improving accountability policies.

Your participation in this interview is valuable as it will shed light on the successes and challenges faced by SEAs in adopting innovative indicators, dual accountability systems, and enhanced local control of accountability. We will specifically discuss your insights on the characteristics of your accountability system, the impact of the flexibilities you have implemented, and any recommendations you may have for future policy improvements.

Confidentiality

Your responses will be treated with utmost confidentiality. In reporting findings, WestEd will not identify you or your organization by name or any other personal characteristics. Participation in this interview is entirely voluntary, and you may decline to answer any questions or withdraw from the study at any time without any repercussions.

Recording and Rights

To ensure accurate documentation of our discussion, I would like to audio record this conversation. The recording will be used solely for research purposes within WestEd and will be securely handled and erased upon completion of the study. If you have any questions regarding your rights as a participant, or if you wish to discuss this further, please feel free to contact _____.

Consent

May I have your permission to record this interview?

[If yes, begin recording. If no, take notes of responses.]

[If yes to the recording question] Could you please state your name and confirm your consent to be recorded?

Thank You.

Thank you for considering this opportunity to contribute to our research. Your insights are invaluable, and we greatly appreciate your time and effort in sharing your experiences with us.

Interview Questions
General Understanding
Can you describe your role within the State Educational Agency and how it relates to implementing ESSA and state accountability?
How has your SEA leveraged ESSA's accountability flexibilities?
Innovative Indicators (if state implemented innovative indicators)
What innovative indicators, if any, has your SEA interwoven into the accountability system under ESSA, beyond standardized test scores?
Dual Accountability Systems (if state implemented dual accountability systems)
How does your SEA manage dual accountability systems balancing both state and federal requirements?
Have there been any specific benefits or disadvantages you've encountered with this dual system approach?

Interview Questions
Enhanced Local Control (if state implemented enhanced local control practices)
In what ways has enhanced local control over accountability practices changed the dynamic between your SEA and Local Educational Agencies?
Can you share any instances where local governance led to successful accountability strategy implementations?

History of Use in the Field (ask all states)

History of innovative practice - how did it come to be?

Impact (ask all states)

What impact has this innovative practice had?

Implementation Strengths & Challenges (ask all states)

What have been the strengths and challenges of implementing this innovative practice?

Are there any additional resources we should be aware of to help us better understand this innovation and your state's implementation?

What else are we missing? Who else do we need to talk to?

WestEd
WestEd.org

Suggested citation: Herz, M., & Wylen, M. (2025). *Refining statewide school accountability systems: Lessons from innovative state approaches under ESSA*. WestEd.

WestEd is a nonpartisan, nonprofit organization that aims to improve the lives of children and adults at all ages of learning and development. We do this by addressing challenges in education and human development, increasing opportunity, and helping build communities where all can thrive. WestEd staff conduct and apply research, provide technical assistance, and support professional learning. We work with early learning educators, classroom teachers, local and state leaders, and policymakers at all levels of government. For more information, visit WestEd.org. For regular updates on research, free resources, solutions, and job postings from WestEd, subscribe to the E-Bulletin, our semimonthly e-newsletter, at WestEd.org/subscribe.